

Concept Presentation

Sixth Open Forum on Metadata Registries

To be held January 20-24, 2003

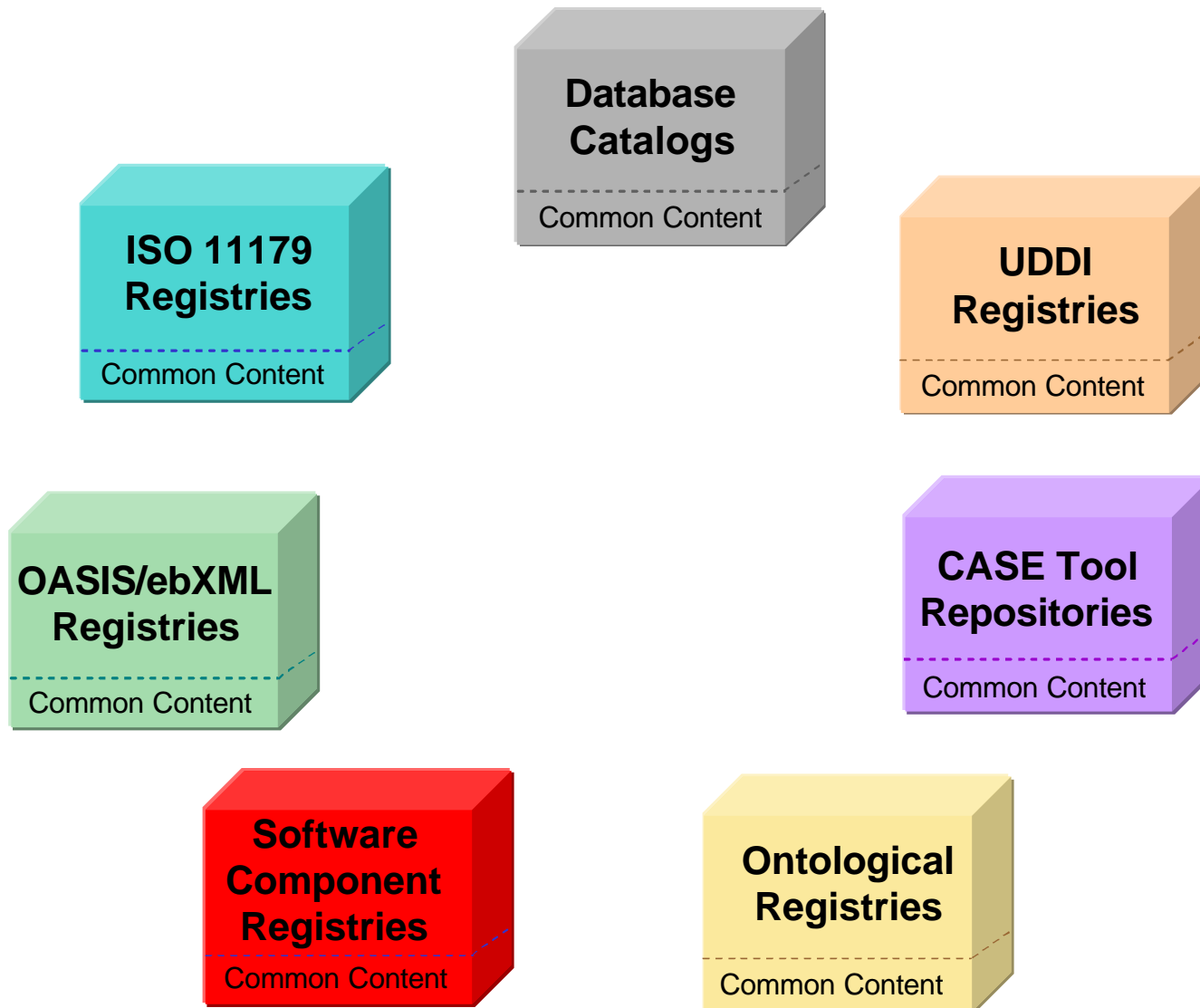
Bruce Bargmeyer
510-495-2905
BEBargmeyer@lbl.gov

Registries



- ➡ Using this term loosely, there are several types of registries in the area of data management and interchange.
- ➡ The registries have some common, overlapping content, which is extended and utilized in different ways.
- ➡ The registries vary according to the intended purpose, granularity of contents, the level of semantics management.

Types of Registries



Types of Registries

☞ OASIS/ebXML XML Registries – XML Artifacts

- ◆ Register XML Artifacts - Includes schemas and DTDs. XML schema relate to a paper form, EDI document, or reporting requirement.
- ◆ Emphasis on syntactic (structure) information.

☞ ISO 11179 Metadata Registries – Data Semantics

- ◆ Register Data Elements, components of data elements and groups of data elements. For example, country codes for customer place of residence. Includes: data element concepts, data elements (including representation), value domains, and (multiple) taxonomies.
- ◆ Emphasis on semantic information such as definitions of data elements and value meanings, and stewardship responsibilities.

☞ Universal Description, Discovery, and Integration (UDDI) Registries – Web-based Business Services

- ◆ A White page directory of the business activities of firms.
- ◆ Emphasis on interface specifications.

Types of Registries (continued)

- ☞ Database System Registries (System Catalogs/Data Dictionaries/ Repositories) – Schema, integrity & operational info.
 - ◆ Includes all elements of database schemas: data elements, relations, integrity constraints.
 - ◆ Emphasis on information required to make database systems work for queries, etc., not on semantic management.

- ☞ Case Tool Registries (Encyclopedias/Repositories) – Data model and application program logic
 - ◆ Includes information needed to create a database and potentially the program code for a system. Contains database schemas.
 - ◆ Emphasis not on semantic management.

- ☞ Ontological Registries – Concept Structures
 - ◆ Includes relations among concepts (subsumption, inheritance,...) and axioms for inference among concepts, e.g., temporal/spatial reasoning, etc.
 - ◆ Emphasis is on semantics.

Types of Registries (continued)

Software Component Registries -- Software Components

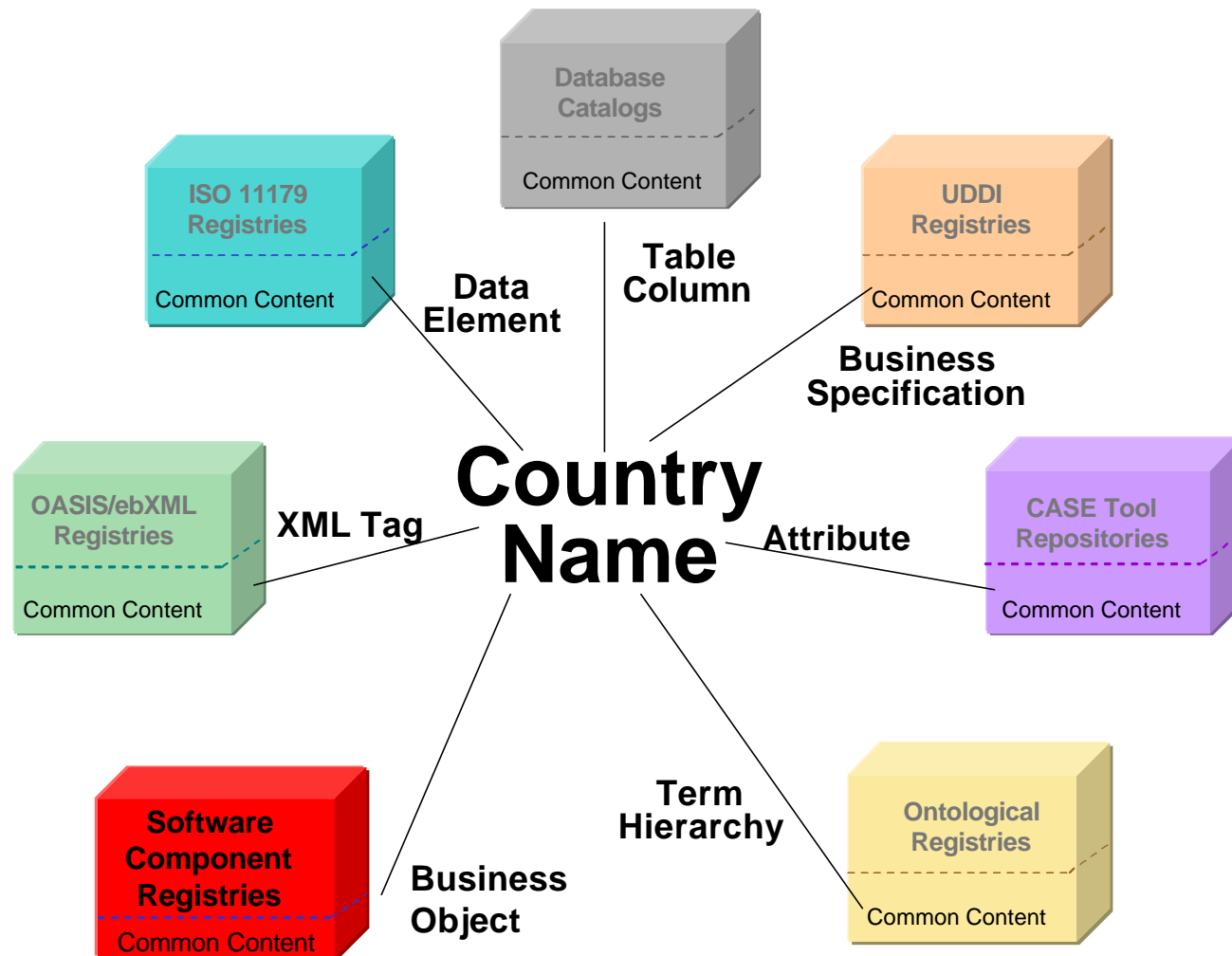
- ◆ Reuse software components built on specific vendor platforms such as EJB, COM, CORBA IDL etc.
- ◆ Include basic common elementary objects and object patterns
- ◆ Emphasis on reusability of software component to build flexible and standardized business systems.

Theme: Cooperation between Registries & Management of Semantics

- ☞ Users work with several of these registries simultaneously, since the registries perform different functions.
- ☞ There is a need for the registries to cooperate (better: to interoperate), since they contain related and sometimes overlapping information. The data must be kept updated and synchronized with a minimum of effort.

Registries

Example of Common Content



Semantics is important to interoperation

Example: Country Identifier

Data Element Concept

Name: Country Identifiers

Context:

Definition:

Unique ID: 5769

Conceptual Domain:

Maintenance Org.:

Steward:

Classification:

Registration Authority:

Others

Algeria
Belgium
China
Denmark
Egypt
France
...
Zimbabwe

Data Elements

Name:
Context:
Definition:
Unique ID: 4572
Value Domain:
Maintenance Org.:
Steward:
Classification:
Registration Authority:
Others

Algeria
Belgium
China
Denmark
Egypt
France
...
Zimbabwe

ISO 3166
English Name

L` Algérie
Belgique
Chine
Danemark
Egypte
La France
...
Zimbabwe

ISO 3166
French Name

DZ
BE
CN
DK
EG
FR
...
ZW

ISO 3166
2-Alpha Code

DZA
BEL
CHN
DNK
EGY
FRA
...
ZWE

ISO 3166
3-Alpha Code

012
056
156
208
818
250
...
716

ISO 3166
3-Numeric Code

ISO 11179 Registry

Data Element Advanced : - Netscape
File Edit View Go Communicator Help

Just Browsing
Standards Implementer
Data Harmonizer
System Developer

Related Programs
EDSC
XML
State Cleanup
One Stop (REI)
EMPACT
EIMS
Facility Information

Utilities
What's New
FAQs
Site Map
Help
Comments
Administration

Related Applications
EDR
TRS
SRS
CRS

Data Element List – Address Group

Name	Definition	Datatype	Status	Compare?
Building Name 1-21293:1	The name of a permanent	Alphanumeric	Incomplete/Awaiting	<input type="checkbox"/>
Building Number 1-5777:1				
City Name 1-6175:1				
Country Code 1-5232:1				
Country Name 1-9:1				
County Name 1-12:1				
Highway Cont Text 1-21296:1				
International P Code 1-5392:1				
Organization Name 1-21291:1				
Person Given Name	The first name and the middle			

Mailing Address: Data Element List

Click on a name to obtain detailed information. The optionality applies to usage in the information source (the optionality given in EPA Data Standards applies only to EPA application systems).

Registry Name	Registry Definition	Type	Optionality	Identifier	Status
Mailing Address City Name	The name delivered.				
Mailing Address Person Given Name	The first name in a				
Mailing Address Person Last Name	The surname				
Mailing Address Person Name Prefix Text	The text the name in a				
Mailing Address State Code	The alpha that repres				
Mailing Address ZIP Code	The code t Improve delivered.				
Mailing Address Highway Contract Route Text	The text th box numbe delivered t				
Mailing Address Post Office Box Number	The numbe addressee				
Mailing Address Rural Route Text	The text th where mail				
Mailing Address Street Text	The text th number wh				
Building Number	The numbe the street t delivery.				
Mailing Address Building Name	The name delivered t				

<?xml version="1.0"?>

<shipTo >

<name>Alice Wilson</name>

<street>161 North Street</street>

<city>Happy Valley</city>

<state>MO</state>

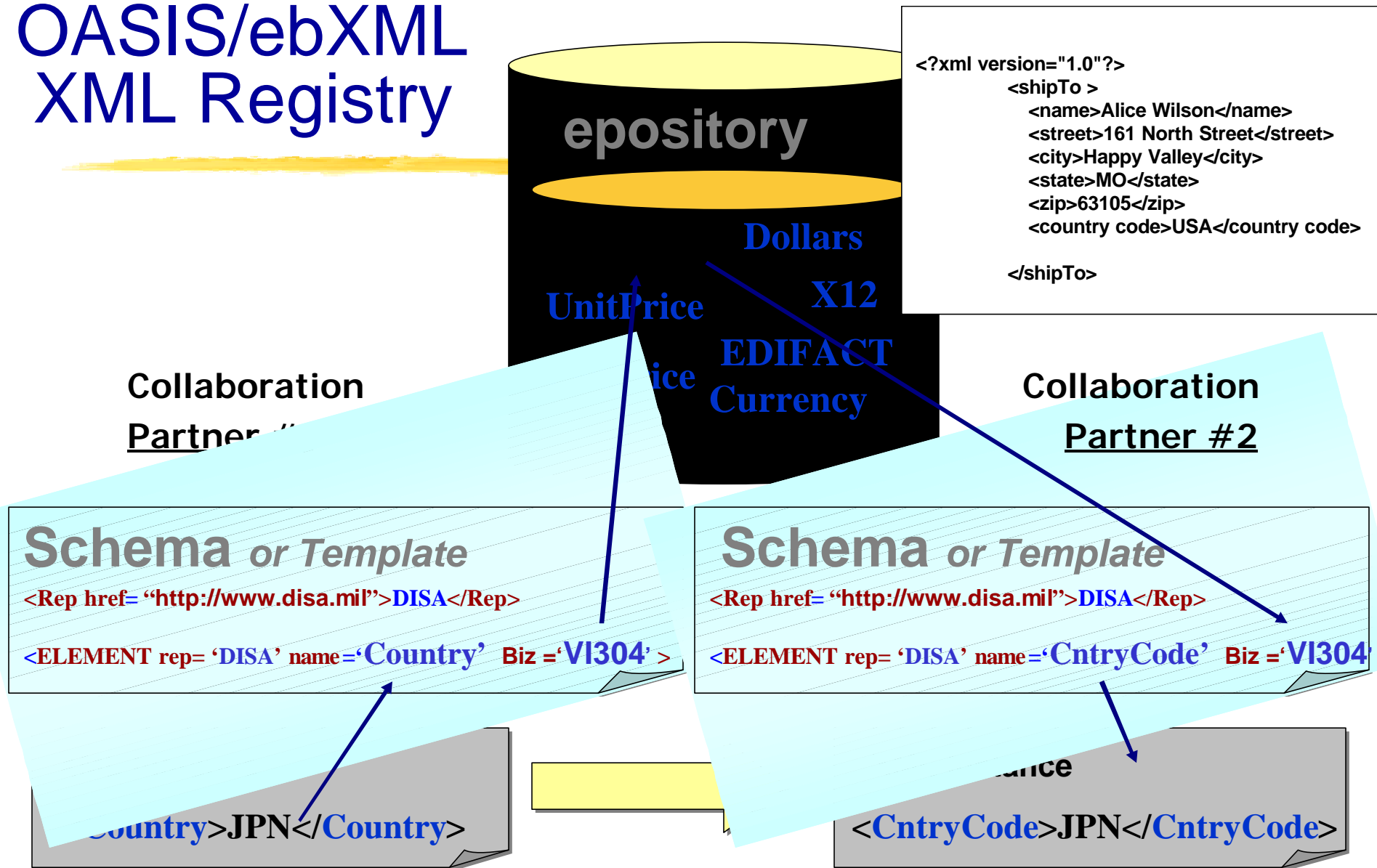
<zip>63105</zip>

<country code>USA</country code>

</shipTo>

Name
Street Address
City, State Postal Code
Country

OASIS/ebXML XML Registry



UDDI Registry



```
<?xml version="1.0"?>
  <shipTo>
    <name>Alice Wilson</name>
    <street>161 North Street</street>
    <city>Happy Valley</city>
    <state>MO</state>
    <zip>63105</zip>
    <country code>USA</country code>
  </shipTo>
```


Database Catalog

The screenshot displays the TOAD (Tool for Oracle and dBase) interface. The main window, titled 'FREE TOAD - [Schema Browser for DMFNEW@DEV]', shows a tree view on the left with the 'CONTACT' table selected. The right pane shows the 'Columns' tab for the 'CONTACT' table, listing columns such as 'CONTACT_ID', 'FIRST_NAME', 'LAST_NAME', 'CREATION_DATE', 'CREATION_USER', 'MIDDLE_INITIAL', 'CONTACT_TITLE', 'STREET_ADDRESS', 'CITY', 'STATE_PROVINCE', 'ZIP_POSTAL_CODE', 'MAIL_CODE', 'COUNTRY', 'PHONE_NUMBER', 'CELL_PHONE_NUMBER', 'PAGER_NUMBER', 'FAX_NUMBER', 'TELEX_NUMBER', 'ELECTRONIC_MAIL_ADDRESS', 'LAST_CHANGE_DATE', and 'LAST_CHANGE_USER'. A 'ddl.tab - Notepad' window is open in the foreground, displaying a SQL script to create a table named 'ADDRESS' with columns: 'STREET_ADDRESS VARCHAR2(50)', 'CITY VARCHAR2(65)', 'STATE_PROVINCE VARCHAR2(100)', 'ZIP_POSTAL_CODE VARCHAR2(10)', 'MAIL_CODE VARCHAR2(15)', and 'COUNTRY VARCHAR2(65)'.

DMFNEW

Indexes Triggers Constraints Seqs
Tables Views Syns Procs

Columns Indexes Constraints Triggers Data Scripts Grants Partitions Stats/Size Referential

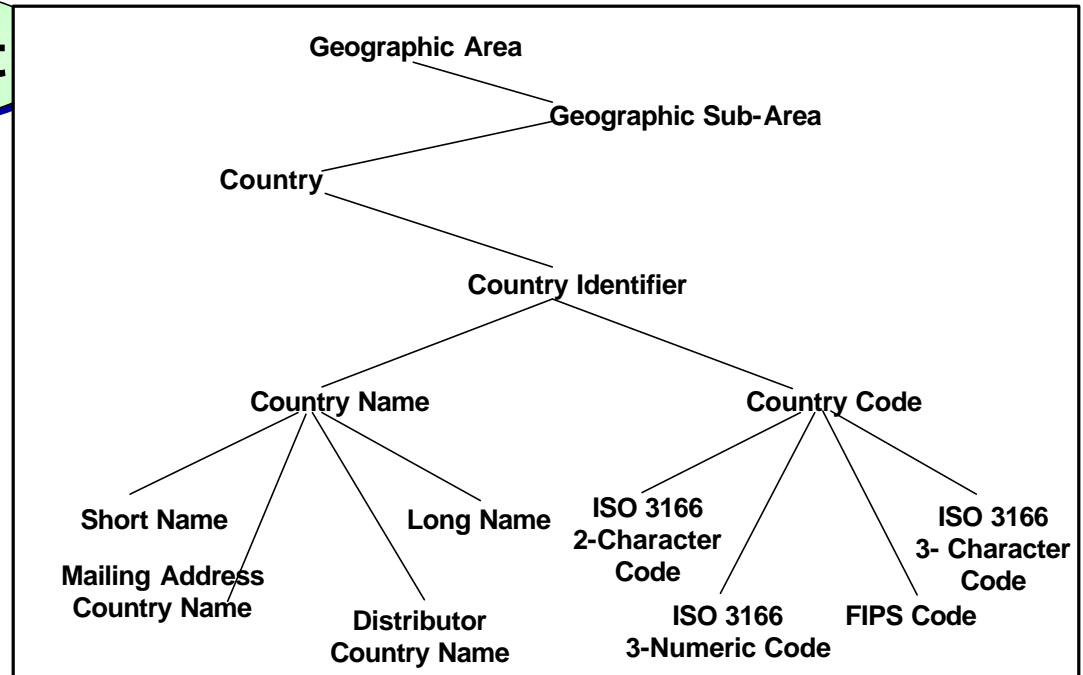
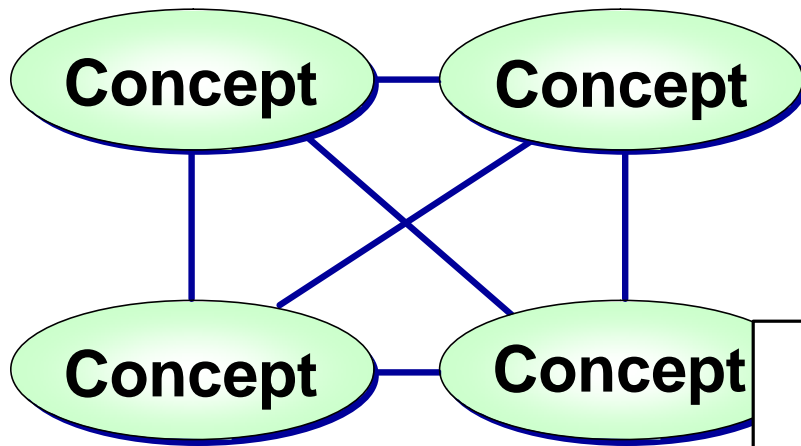
Show no comments CONTACT_ID

Column	Data Type	Null?	Default
CONTACT_ID	NUMBER (9)	N	
FIRST_NAME	VARCHAR2 (200)	N	
LAST_NAME	VARCHAR2 (200)	N	
CREATION_DATE	DATE	N	
CREATION_USER	VARCHAR2 (30)	N	
MIDDLE_INITIAL	VARCHAR2 (10)	Y	
CONTACT_TITLE	VARCHAR2 (30)	Y	
STREET_ADDRESS	VARCHAR2 (50)	Y	
CITY	VARCHAR2 (65)	Y	
STATE_PROVINCE	VARCHAR2 (100)	Y	
ZIP_POSTAL_CODE	VARCHAR2 (10)	Y	
MAIL_CODE	VARCHAR2 (15)	Y	
COUNTRY	VARCHAR2 (65)	Y	
PHONE_NUMBER	VARCHAR2 (15)	Y	
CELL_PHONE_NUMBER	VARCHAR2 (15)	Y	
PAGER_NUMBER	VARCHAR2 (15)	Y	
FAX_NUMBER	VARCHAR2 (15)	Y	
TELEX_NUMBER	VARCHAR2 (15)	Y	
ELECTRONIC_MAIL_AD...	VARCHAR2 (100)	Y	
LAST_CHANGE_DATE	DATE	Y	
LAST_CHANGE_USER	VARCHAR2 (30)	Y	

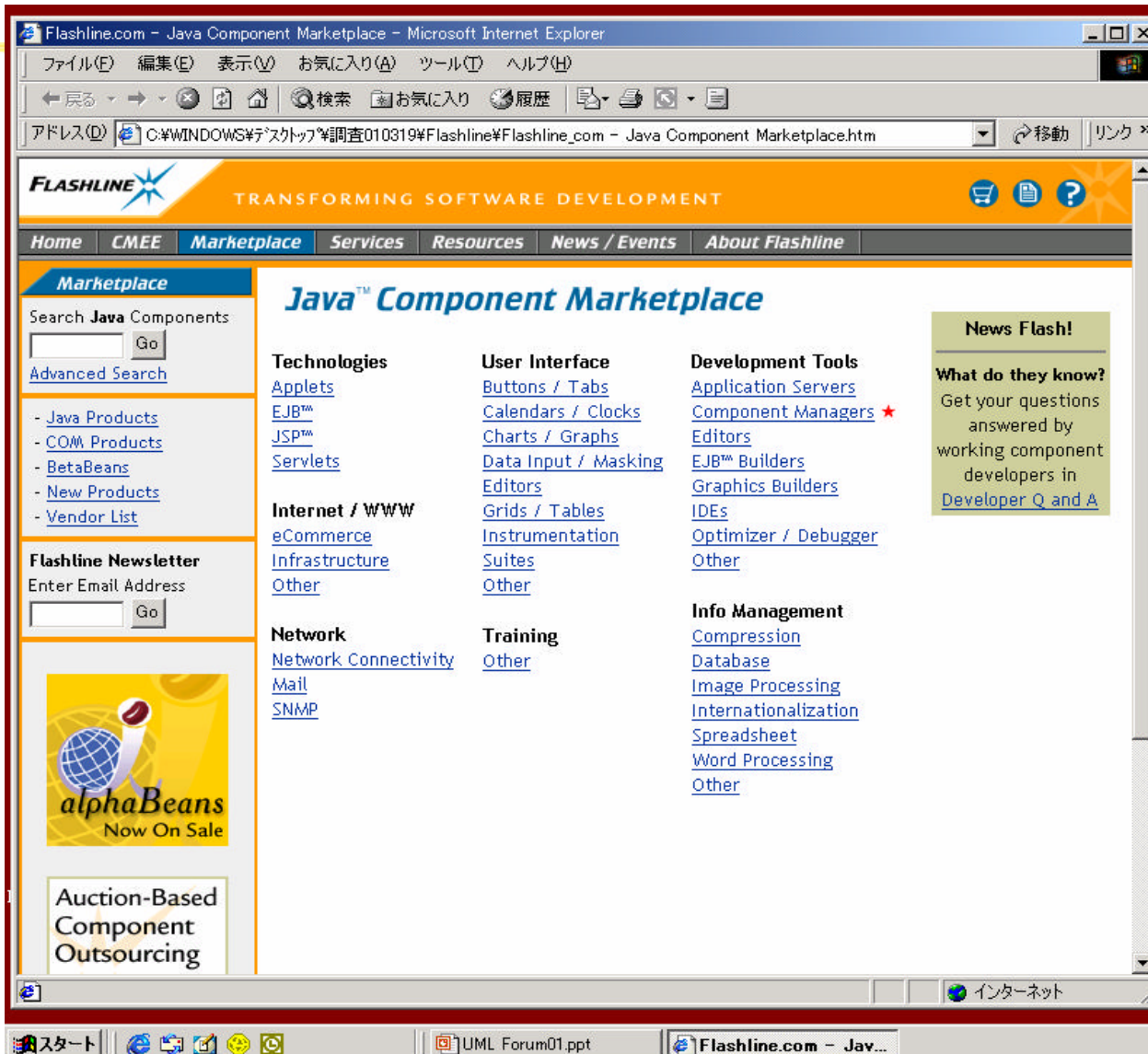
ddl.tab - Notepad

```
PROMPT Creating Table 'ADDRESS'
CREATE TABLE ADDRESS
(STREET_ADDRESS VARCHAR2(50)
,CITY VARCHAR2(65)
,STATE_PROVINCE VARCHAR2(100)
,ZIP_POSTAL_CODE VARCHAR2(10)
,MAIL_CODE VARCHAR2(15)
,COUNTRY VARCHAR2(65))
```

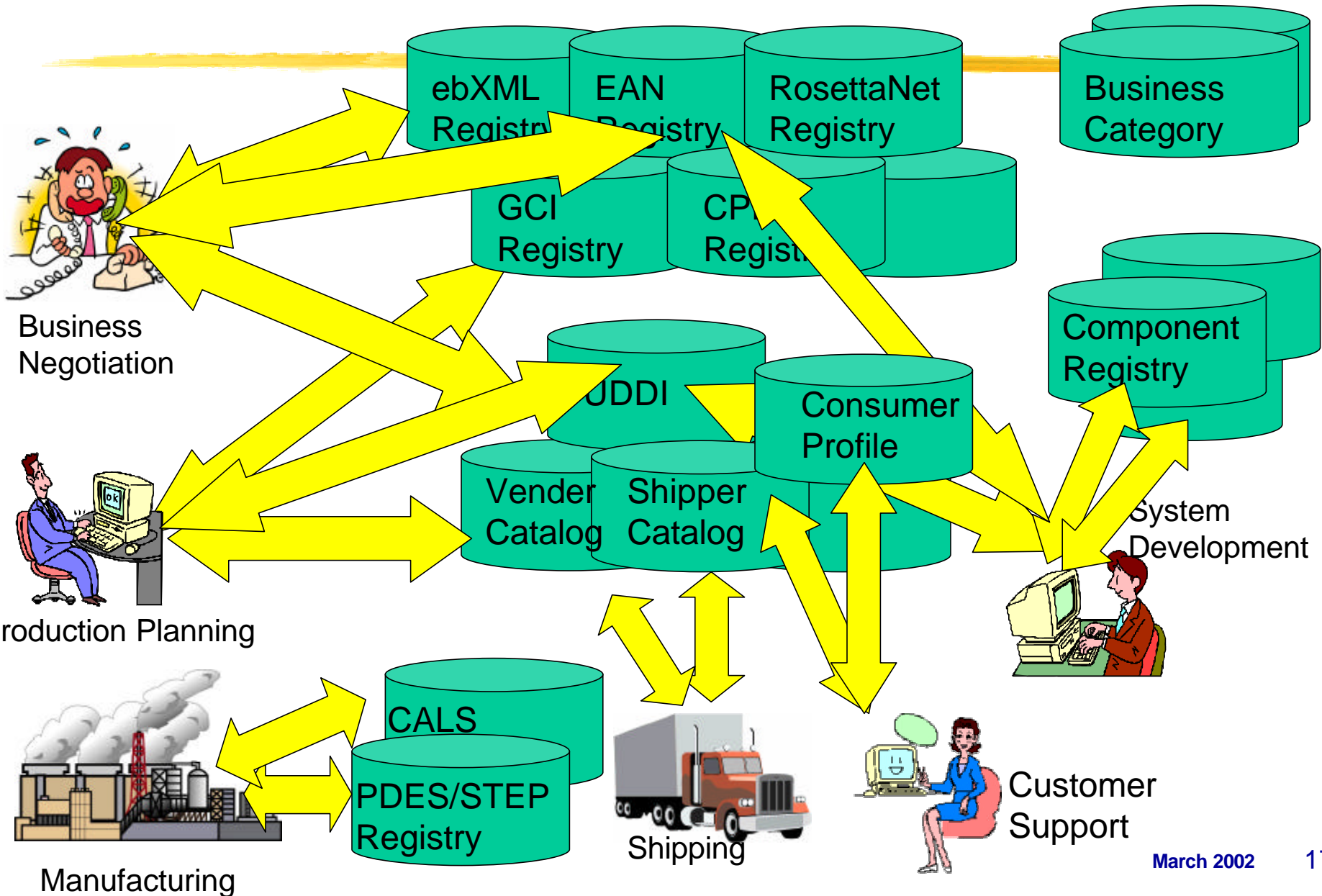
Ontological Registry



An Example of Software Component Registry



Needs for Registry Interoperation

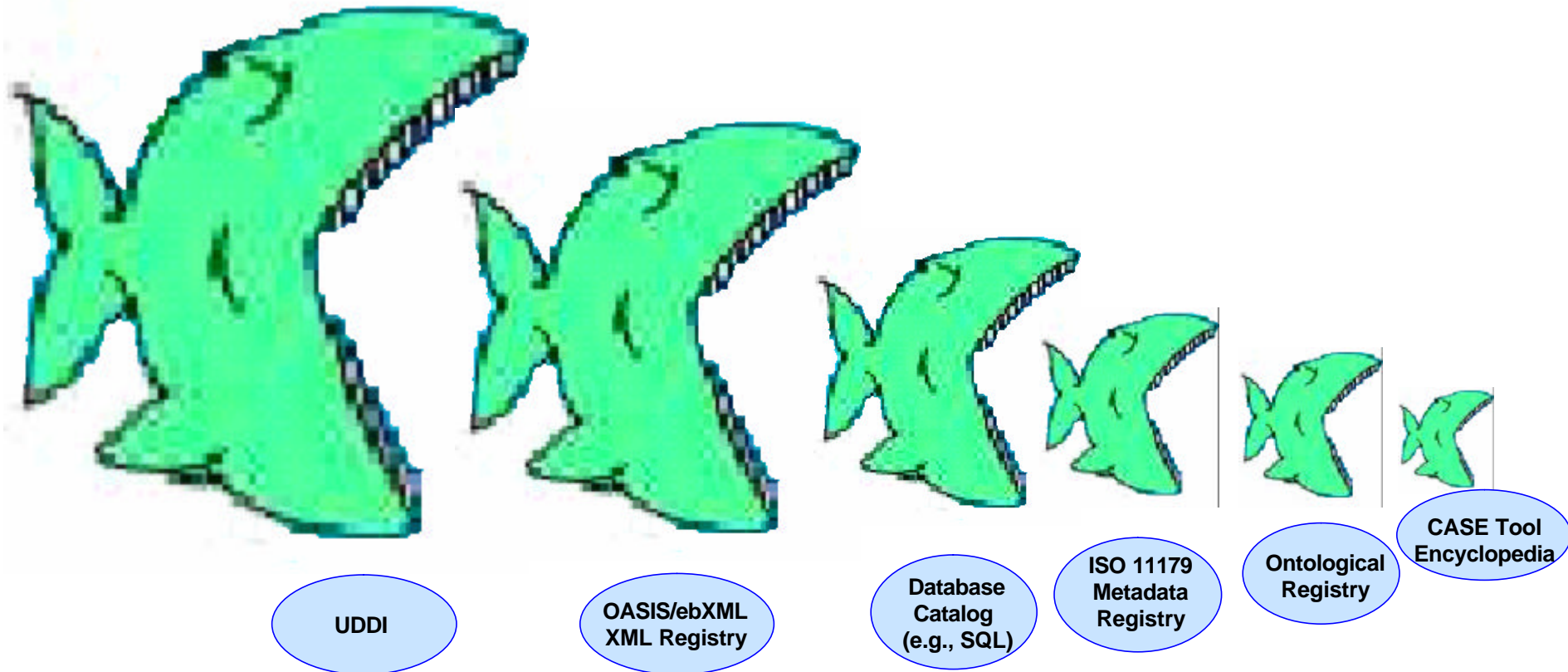


Consolidation or Interoperation?

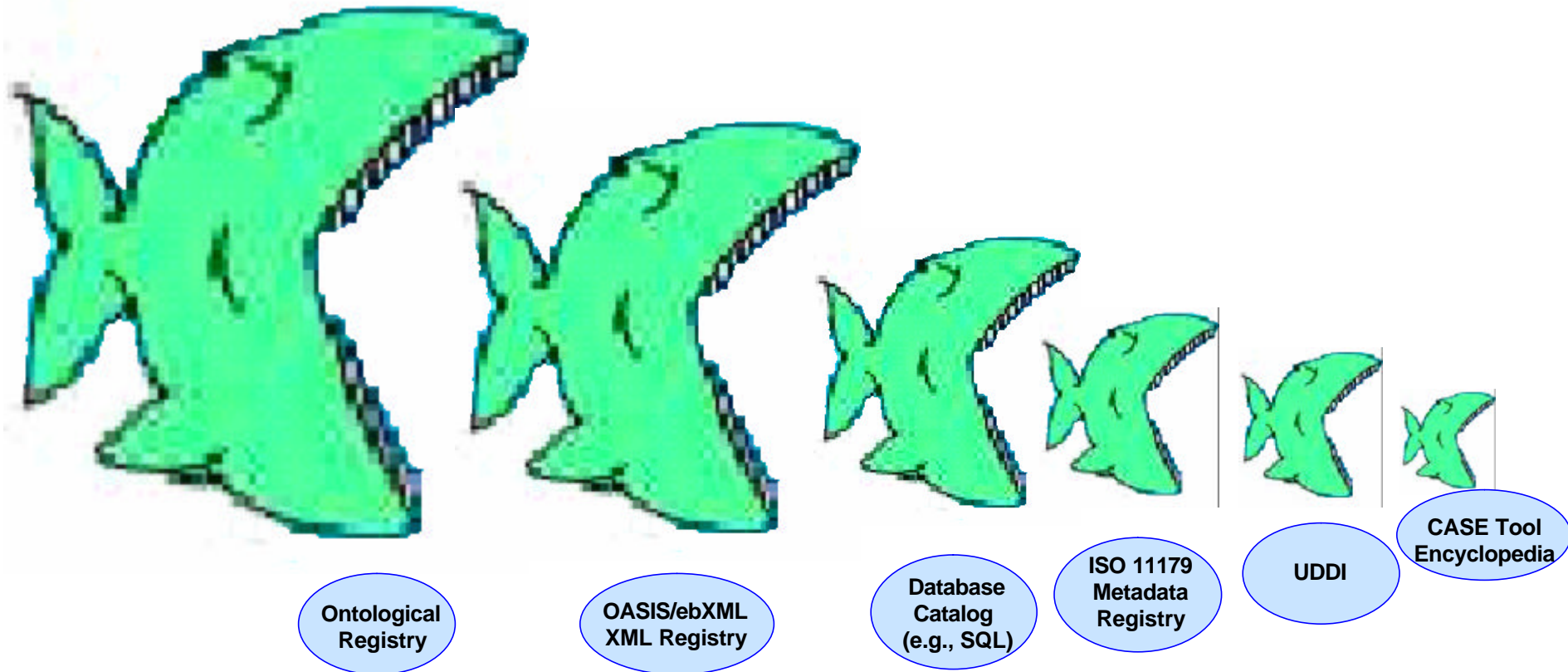
(continued)

- ☞ Does one registry acquire all of the key functionalities?
- ☞ Do distinct capabilities require multiple registries with interoperability?

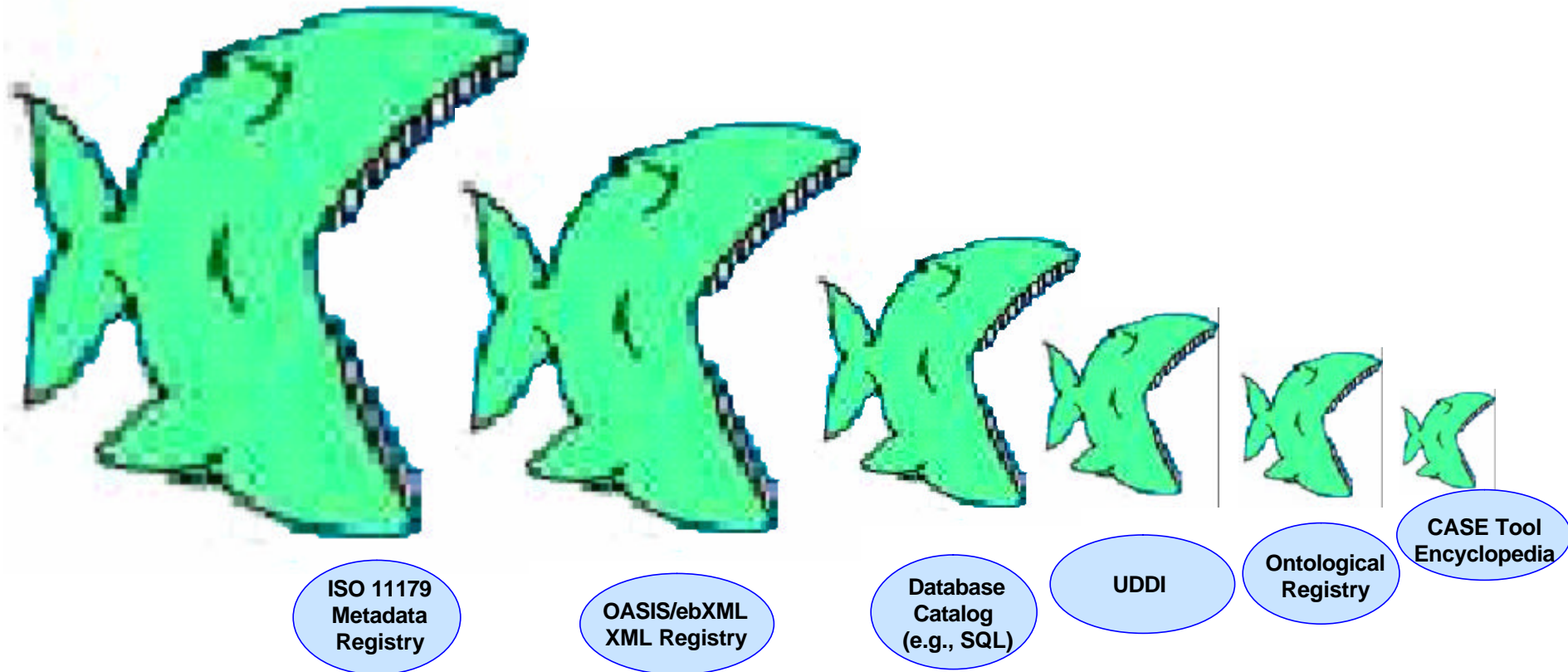
Consolidation or Interoperation?



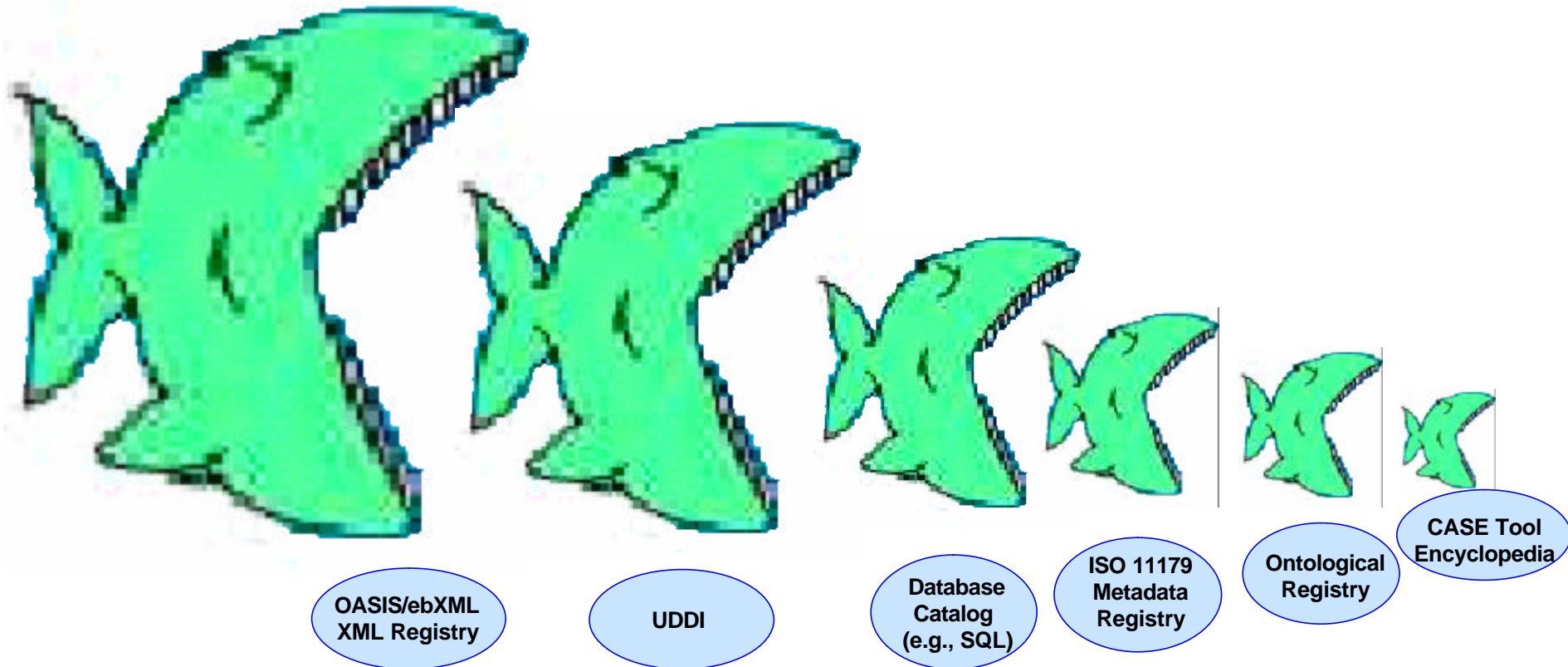
Consolidation or Interoperation?



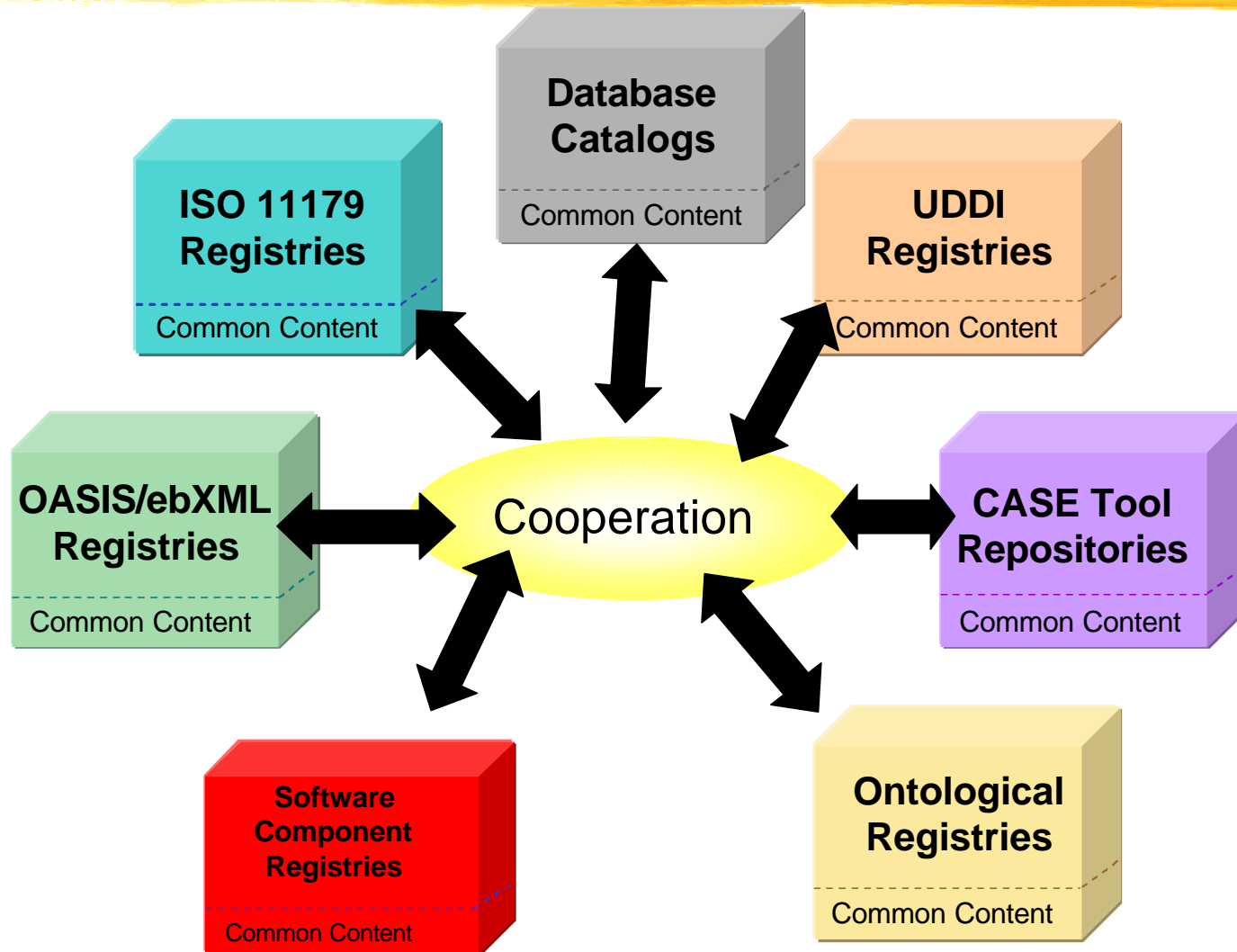
Consolidation or Interoperation?



Consolidation or Interoperation?



Registry types



Open Forum on Metadata Registries

Semantic Interoperability between Registries

Major topic:

- ◆ How to implement and use registries,
- ◆ cooperation between registries
- ◆ semantics management

Participants:

- ◆ Standards developers, Software developers, Practitioners.

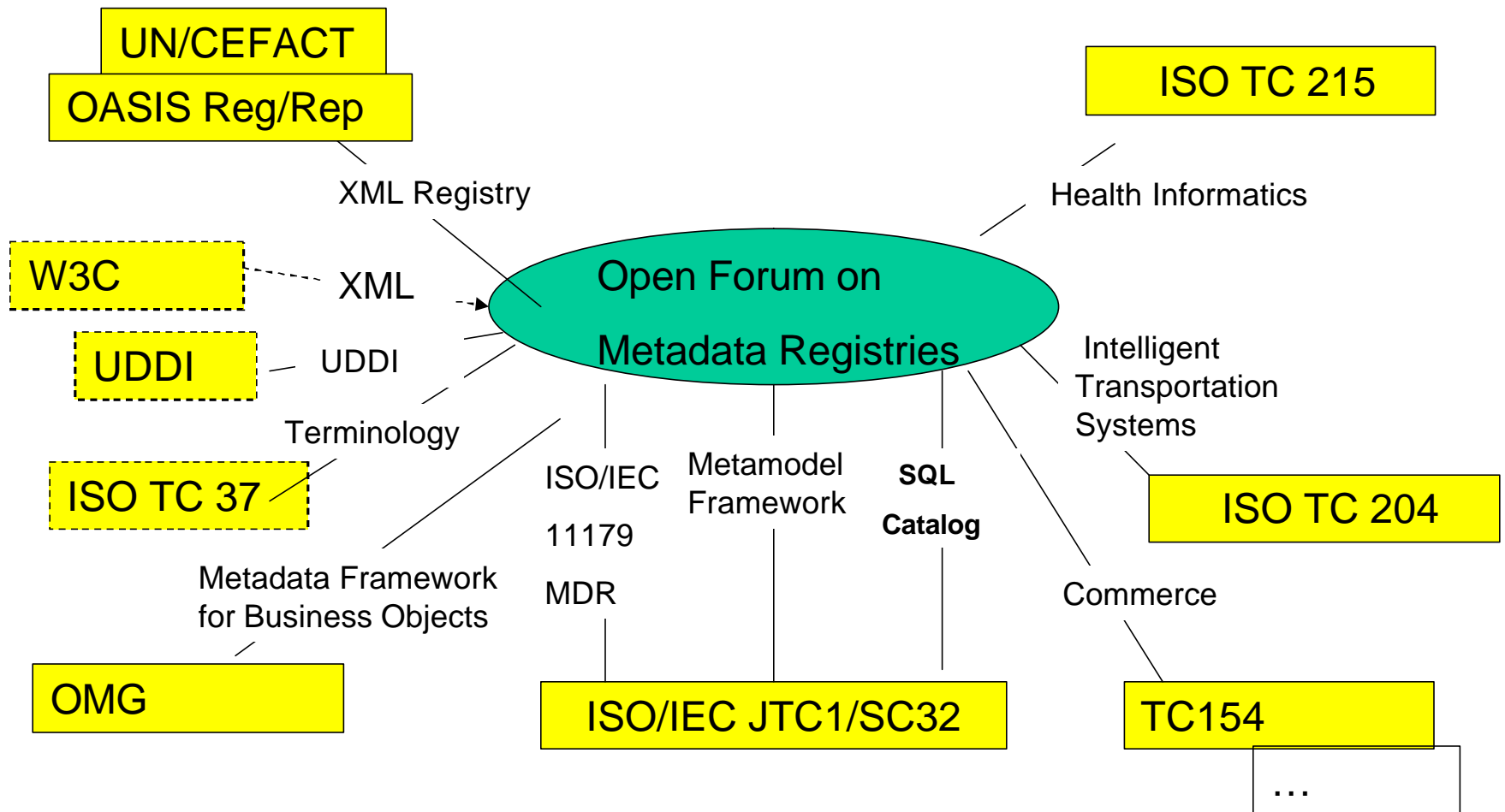
Related topics:

- ◆ Tutorials on the standards.
- ◆ Tracks with demonstrations for application areas

Conference is intended to describe the registries, demonstrate their use, encourage progress toward cooperation between registries, and showcase progress made.

Some Standards Development Organizations

(Related to Registries)



Tracks – Standards and Working Implementations

Standards:

- ➡ Standards Track

Practitioner Tracks:

- ➡ Defense Track
- ➡ Environment Track
- ➡ Healthcare Track
- ➡ Learning Technologies and Knowledge Management Track
- ➡ Statistics Track
- ➡ Transportation & Aviation Track
- ➡ Electronic Business Track?

In Sum:

- ➡ A conference drawing together standards developers, software developers and practitioners.
- ➡ The conference is intended to introduce the registries, show how the registries are used and describe the related standards. A major topic will be cooperation between the registries to manage semantics.
- ➡ ISO/IEC 11179 is primarily aimed at semantics management and will be the focus of demonstrations in some of the practitioner tracks. Special emphasis will be given to interoperability between 11179 registries, UDDI registries, XML registries/repositories, database catalogs, etc.
- ➡ The conference organizers will work with standards developers, software developers and practitioners to encourage progress toward cooperation & interoperability. The conference is intended to showcase progress made.
- ➡ The conference will have separate tracks for various communities of interest.
- ➡ Presentations are by invitation
- ➡ The conference will have multiple sponsors/hosts, with ISO/IEC JTC 1 SC32 (Data Management and Interchange) /WG 2 (Metadata) as the initial sponsor and EPA, the Federal CIO Council's XML Working Group and the Lawrence Berkeley National Laboratory as the initial hosts.